#4/1
EWallace
7-29-02

Atty Dkt No. 8325-0011.21
S11-US3
09/990,186

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Assistant Commissioner for Patents, Washington, D.C. 20231 on
05 - FEBRUARY - 2002.

2/5/02
Date

Signature

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Application of:

LIU

Confirmation No.: 1799

Serial No.: 09/990,186

Group Art Unit: Unassigned

Filing Date: November 20, 2001

Examiner: Unassigned

Title: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE TRIPLETS BY ZINC FINGERS

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

This preliminary amendment is filed prior to substantive examination in the above-referenced case. No fee is believed to be due in connection with the filing of this amendment. Consideration of the following amendments is requested prior to substantive examination.

I. AMENDMENTS

In the specification:

Please replace the paragraph beginning on page 4, line 13 with the following rewritten paragraph:

Thus, provided herein is a zinc finger protein that binds to a target site, said zinc finger protein comprising a first (F1), a second (F2), and a third (F3) zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, said target site comprising, in 3' to 5' direction, a first (S1), a second (S2), and a third (S3) target subsite, each target subsite having the nucleotide sequence GNN, wherein if S1 comprises GAA, F1 comprises the amino acid sequence QRSNLVR (SEQ ID NO:158); if S2 comprises GAA, F2 comprises the amino acid sequence QSGNLAR (SEQ ID NO:801); if S3 comprises GAA, F3 comprises the amino acid sequence QSGNLAR (SEQ ID NO:801); if S1 comprises GAG, F1 comprises the amino acid sequence RSDNLAR (SEQ ID NO:130); if S2 comprises GAG, F2 comprises the amino acid sequence RSDNLAR (SEQ ID NO:130); if S3 comprises GAG, F3 comprises the amino acid sequence RSDNLTR (SEQ ID NO:231); if S1 comprises GAC, F1 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S2 comprises GAC, F2 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S3 comprises GAC, F3 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S1 comprises GAT, F1 comprises the amino acid sequence QSSNLAR (SEQ ID NO:1765); if S2 comprises GAT, F2 comprises the amino acid sequence TSGNLVR (SEQ ID NO:1442); if S3 comprises GAT, F3 comprises the amino acid sequence TSANLSR (SEQ ID NO:377); if S1 comprises GGA, F1 comprises the amino acid sequence QSGHLAR (SEQ ID NO:413); if S2 comprises GGA, F2 comprises the amino acid sequence QSGHLQR (SEQ ID NO:287); if S3 comprises GGA, F3 comprises the amino acid sequence QSGHLQR (SEQ ID NO:287); if S1 comprises GGG, F1 comprises the amino acid sequence RSDHLAR (SEQ ID NO:127); if S2 comprises GGG, F2 comprises the amino acid sequence RSDHLSR (SEQ ID NO:229); if S3 comprises GGG, F3 comprises the amino acid sequence RSDHLSR (SEQ ID

NO:229); if S1 comprises GGC, F1 comprises the amino acid sequence DRSHLRT (SEQ ID NO:1506); if S2 comprises GGC, F2 comprises the amino acid sequence DRSHLAR (SEQ ID NO:1092); if S1 comprises GGT, F1 comprises the amino acid sequence QSSHLTR (SEQ ID NO:835); if S2 comprises GGT, F2 comprises the amino acid sequence TSGHLSR (SEQ ID NO:1201); if S3 comprises GGT, F3 comprises the amino acid sequence TSGHLVR (SEQ ID NO:1425); if S1 comprises GCA, F1 comprises the amino acid sequence QSGSLTR (SEQ ID NO:342); if S2 comprises GCA, F2 comprises QSGDLTR (SEQ ID NO:220); if S3 comprises GCA, F3 comprises QSGDLTR (SEQ ID NO:220); if S1 comprises GCG, F1 comprises the amino acid sequence RSDDLTR (SEQ ID NO:188); if S2 comprises GCG, F2 comprises the amino acid sequence RSDDLQR (SEQ ID NO:1844); if S3 comprises GCG, F3 comprises the amino acid sequence RSDDLTR (SEQ ID NO:188); if S1 comprises GCC, F1 comprises the amino acid sequence ERGTLAR (SEQ ID NO:131); if S2 comprises GCC, F2 comprises the amino acid sequence DRSDLTR (SEQ ID NO:417); if S3 comprises GCC, F3 comprises the amino acid sequence DRSDLTR (SEQ ID NO:417); if S1 comprises GCT, F1 comprises the amino acid sequence QSSDLTR (SEQ ID NO:1450); if S2 comprises GCT, F2 comprises the amino acid sequence QSSDLTR (SEQ ID NO:1450); if S3 comprises GCT, F3 comprises the amino acid sequence QSSDLQR (SEQ ID NO:132); if S1 comprises GTA, F1 comprises the amino acid sequence QSGALTR (SEQ ID NO:1398); if S2 comprises GTA, F2 comprises the amino acid sequence QSGALAR (SEQ ID NO:3339); if S1 comprises GTG, F1 comprises the amino acid sequence RSDALTR (SEQ ID NO:153); if S2 comprises GTG, F2 comprises the amino acid sequence RSDALSR (SEQ ID NO:237); if S3 comprises GTG, F3 comprises the amino acid sequence RSDALTR (SEQ ID NO:153); if S1 comprises GTC, F1 comprises the amino acid sequence DRSALAR (SEQ ID NO:184); if S2 comprises GTC, F2 comprises the amino acid sequence DRSALAR (SEQ ID NO:184); and if S3 comprises GTC, F3 comprises the amino acid sequence DRSALAR (SEQ ID NO:184).

Please replace the paragraph beginning on page 5, line 23 with the following rewritten

paragraph:

[47] Also provided are methods of designing a zinc finger protein comprising a first (F1), a second (F2), and a third (F3) zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus that binds to a target site comprising, in 3' to 5' direction, a first (S1), a second (S2), and a third (S3) target subsite, each target subsite having the nucleotide sequence GNN, the method comprising the steps of (a) selecting the F1 zinc finger such that it binds to the S1 target subsite, wherein if S1 comprises GAA, F1 comprises the amino acid sequence QRSNLVR (SEQ ID NO:158); if S1 comprises GAG, F1 comprises the amino acid sequence RSDNLAR (SEQ ID NO:130); if S1 comprises GAC, F1 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S1 comprises GAT, F1 comprises the amino acid sequence QSSNLAR (SEQ ID NO:1765); if S1 comprises GGA, F1 comprises the amino acid sequence QSGHLAR (SEQ ID NO:413); if S1 comprises GGG, F1 comprises the amino acid sequence RSDHLAR (SEQ ID NO:127); if S1 comprises GGC, F1 comprises the amino acid sequence DRSHLRT (SEQ ID NO:1506); if S1 comprises GGT, F1 comprises the amino acid sequence QSSHLTR (SEQ ID NO:835); if S1 comprises GCA, F1 comprises QSGSLTR (SEQ ID NO:342); if S1 comprises GCG, F1 comprises RSDDLTR (SEQ ID NO:188); if S2 comprises GCG, F2 comprises RSDDLQR (SEQ ID NO:1844); if S1 comprises GCC, F1 comprises ERGTLAR (SEQ ID NO:131); if S1 comprises GCT, F1 comprises the amino acid sequence QSSDLTR (SEQ ID NO:1450); if S1 comprises GTA, F1 comprises the amino acid sequence QSGALTR (SEQ ID NO:1398); if S1 comprises GTG, F1 comprises the amino acid sequence RSDALTR (SEQ ID NO:153); if S1 comprises GTC, F1 comprises the amino acid sequence DRSALAR (SEQ ID NO:184); (b) selecting the F2 zinc finger such that it binds to the S2 target subsite, wherein S2 comprises GAA, F2 comprises the amino acid sequence QSGNLAR (SEQ ID NO:801); if S2 comprises GAG, F2 comprises the amino acid sequence RSDNLAR (SEQ ID NO:130); if S2 comprises GAC, F2 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S2 comprises GAT, F2 comprises the amino acid sequence TSGNLVR (SEQ ID NO:1442); if S2 comprises GGA, F2 comprises the amino acid sequence QSGHLQR (SEQ ID NO:287); if S2 comprises

GGG, F2 comprises the amino acid sequence RSDHLSR (SEQ ID NO:229); if S2 comprises GGC, F2 comprises the amino acid sequence DRSHLAR (SEQ ID NO:1092); if S2 comprises GGT, F2 comprises the amino acid sequence TSGHLSR (SEQ ID NO:1201); if S2 comprises GCA, F2 comprises the amino acid sequence QSGDLTR (SEQ ID NO:220); if S2 comprises GCC, F2 comprises the amino acid sequence DRSDLTR (SEQ ID NO:417); if S2 comprises GCT, F2 comprises the amino acid sequence QSSDLTR (SEQ ID NO:1450); if S2 comprises GTA, F2 comprises the amino acid sequence QSGALAR (SEQ ID NO:3339); if S2 comprises GTG, F2 comprises the amino acid sequence RSDALSR (SEQ ID NO:237); if S2 comprises GTC, F2 comprises the amino acid sequence DRSALAR (SEQ ID NO:184); and (c) selecting the F3 zinc finger such that it binds to the S3 target subsite, wherein if S3 comprises GAA, F3 comprises the amino acid sequence QSGNLAR (SEQ ID NO:801); if S3 comprises GAG, F3 comprises the amino acid sequence RSDNLTR (SEQ ID NO:231); if S3 comprises GAC, F3 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S3 comprises GAT, F3 comprises the amino acid sequence TSANLSR (SEQ ID NO:377); if S3 comprises GGA, F3 comprises the amino acid sequence QSGHLQR (SEQ ID NO:287); if S3 comprises GGG, F3 comprises RSDHLSR (SEQ ID NO:229); if S3 comprises GGT, F3 comprises the amino acid sequence TSGHLVR (SEQ ID NO:1425); if S3 comprises GCA, F3 comprises the amino acid sequence QSGDLTR (SEQ ID NO:220); if S3 comprises GCG, F3 comprises the amino acid sequence RSDDLTR (SEQ ID NO:188); if S3 comprises GCC, F3 comprises the amino acid sequence DRSDLTR (SEQ ID NO:417); if S3 comprises GCT, F3 comprises the amino acid sequence QSSDLQR (SEQ ID NO:132); if S3 comprises GTG, F3 comprises RSDALTR (SEQ ID NO:153); and if S3 comprises GTC, F3 comprises the amino acid sequence DRSALAR (SEQ ID NO:184);

thereby designing a zinc finger protein that binds to a target site.

Please replace the paragraph beginning on page 7, line 5 with the following rewritten paragraph:

In certain embodiments of the zinc finger proteins and methods described herein, S1 comprises GAA and F1 comprises the amino acid sequence QRSNLVR (SEQ ID NO:158). In other embodiments, S2 comprises GAA and F2 comprises the amino acid sequence QSGNLAR (SEQ ID NO:801). In other embodiments, S3 comprises GAA and F3 comprises the amino acid sequence QSGNLAR (SEQ ID NO:801). In other embodiments, S1 comprises GAG and F1 comprises the amino acid sequence RSDNLAR (SEQ ID NO:130). In other embodiments, S2 comprises GAG and F2 comprises the amino acid sequence RSDNLAR (SEQ ID NO:130). In other embodiments, S3 comprises GAG and F3 comprises the amino acid sequence RSDNLTR (SEQ ID NO:231). In other embodiments, S1 comprises GAC and F1 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395). In other embodiments, S2 comprises GAC and F2 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395). In other embodiments, S3 comprises GAC and F3 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395). In other embodiments, S1 comprises GAT and F1 comprises the amino acid sequence QSSNLAR (SEQ ID NO:1765). In other embodiments, S2 comprises GAT and F2 comprises the amino acid sequence TSGNLVR (SEQ ID NO:1442). In other embodiments, S3 comprises GAT and F3 comprises the amino acid sequence TSANLSR (SEQ ID NO:377). In other embodiments, S1 comprises GGA and F1 comprises the amino acid sequence QSGHLAR (SEQ ID NO:413). In other embodiments, S2 comprises GGA and F2 comprises the amino acid sequence QSGHLQR (SEQ ID NO:287). In other embodiments, S3 comprises GGA and F3 comprises the amino acid sequence QSGHLQR (SEQ ID NO:287). In other embodiments, S1 comprises GGG and F1 comprises the amino acid sequence RSDHLAR (SEQ ID NO:127). In other embodiments, S2 comprises GGG and F2 comprises the amino acid sequence RSDHLSR (SEQ ID NO:229). In other embodiments, S3 comprises GGG and F3 comprises the amino acid sequence RSDHLSR (SEQ ID NO:229). In other embodiments, S1 comprises GGC and F1 comprises the amino acid sequence DRSHLTR (SEQ ID NO:705). In other embodiments, S2 comprises GGC and F2 comprises the amino acid sequence DRSHLAR (SEQ ID NO:1092). In other embodiments, S1 comprises GGT and F1 comprises the amino acid sequence QSSHLTR (SEQ ID NO:835). In

A3

other embodiments, S2 comprises GGT and F2 comprises the amino acid sequence TSGHLSR (SEQ ID NO:1201). In other embodiments, S3 comprises GGT and F3 comprises the amino acid sequence TSGHLVR (SEQ ID NO:1425). In other embodiments, S1 comprises GCA and F1 comprises the amino acid sequence QSGSLTR (SEQ ID NO:342). In other embodiments, S2 comprises GCA and F2 comprises the amino acid sequence QSGDLTR (SEQ ID NO:220). In other embodiments, S3 comprises GCA and F3 comprises the amino acid sequence QSGDLTR (SEQ ID NO:220). In other embodiments, S1 comprises GCG and F1 comprises the amino acid sequence RSDDLTR (SEQ ID NO:188). In other embodiments, S2 comprises GCG and F2 comprises the amino acid sequence RSDDLQR (SEQ ID NO:1844). In other embodiments, S3 comprises GCG and F3 comprises the amino acid sequence RSDDLTR (SEQ ID NO:188). In other embodiments, S1 comprises GCC and F1 comprises the amino acid sequence ERGTLAR (SEQ ID NO:131). In other embodiments, S2 comprises GCC and F2 comprises the amino acid sequence DRSDLTR (SEQ ID NO:417). In other embodiments, S3 comprises GCC and F3 comprises the amino acid sequence DRSDLTR (SEQ ID NO:417). In other embodiments, S1 comprises GCT and F1 comprises the amino acid sequence QSSDLTR (SEQ ID NO:1450). In other embodiments, S2 comprises GCT and F2 comprises the amino acid sequence QSSDLTR (SEQ ID NO:1450). In other embodiments, S3 comprises GCT and F3 comprises the amino acid sequence QSSDLQR (SEQ ID NO:132). In other embodiments, S1 comprises GTA and F1 comprises the amino acid sequence QSGALTR (SEQ ID NO:1398). In other embodiments, S2 comprises GTA and F2 comprises the amino acid sequence QSGALAR (SEQ ID NO:3339). In other embodiments, S1 comprises GTG and F1 comprises the amino acid sequence RSDALTR (SEQ ID NO:153). In other embodiments, S2 comprises GTG and F2 comprises the amino acid sequence RSDALSR (SEQ ID NO:237). In other embodiments, S3 comprises GTG and F3 comprises the amino acid sequence RSDALTR (SEQ ID NO:153). In other embodiments, S1 comprises GTC and F1 comprises the amino acid sequence DRSALAR (SEQ ID NO:184). In other embodiments, S2 comprises GTC and F2 comprises the amino acid sequence DRSALAR (SEQ ID NO:184). In other embodiments, S3 comprises GTC and F3 comprises the amino acid

A3 cont'd

A3 Cont'd

sequence DRSALAR (SEQ ID NO:184).

Please replace the paragraph beginning on page 9, line 10 with the following rewritten paragraph:

Figure 1 shows results of site selection analysis of two representative zinc finger proteins (leftmost 4 columns) and measurements of binding affinity for each of these proteins to their intended target sequences and to variant target sequences. (rightmost 3 columns). Analysis of ZFP1 is shown in the upper portion of the figure and analysis of ZFP2 is shown in the lower portion of the figure. For the site selection analyses, the amino acid sequences of residues –1 through +6 of the recognition helix of each of the three component zinc fingers (F3 (ZFP1, SEQ ID NO:130; ZFP2, SEQ ID NO:420), F2 (ZFP1, SEQ ID NO:1051; ZFP2, SEQ ID NO:889) and F1 (ZFP1, SEQ ID NO:395; ZFP2, SEQ ID NO:685)) are shown across the top row; the intended target sequence (divided into finger-specific target subsites) is shown across the second row, and a summary of the sequences bound is shown in the third row. Data for F3 is shown in the second column, data for F2 is shown in the third column, and data for F1 is shown in the third column.

A4

Please replace the paragraph beginning on page 9, line 24 with the following rewritten paragraph:

Figure 2 shows amino acid sequences of zinc finger recognition regions (amino acids –1 through +6 of the recognition helix) that bind to each of the 16 GNN triplet subsites. Three amino acid sequences are shown for each trinucleotide subsite; these correspond to optimal amino acid sequences for recognition of the subsite from each of the three positions (finger 1, F1(SEQ ID NOS:688, 2534, 676, 1769, 342, 1450, 131, 158, 1765, 395, 1407, 2644, 705, 1398, 1733 & 184); finger 2, F2 (SEQ ID NOS:688, 229, 1446, 3051, 220, 1450, 417, 801, 1442, 395, 1824, 1201, 972, 3339, 1151 & 184); or finger 3, F3 (SEQ ID NOS:943, 229, 676, 1769, 220, 1365, 417, 801, 3525, 395, 1824, 1425, 972, 3592, 952 & 184)) in a three-finger zinc finger protein. Amino acid sequences are from N-terminal to C-terminal; nucleotide sequences are

A5

from 5' to 3'.

Please replace the paragraph beginning on page 10, line 1 with the following rewritten paragraph:

Also shown are site selection results for each of the 48 position-dependent GNN-recognizing zinc fingers. These show the number of times a particular nucleotide was present, at a given position, in a collection of oligonucleotide sequences bound by the finger. For example, out of 15 oligonucleotides bound by a zinc finger protein with the amino acid sequence QSGHLAR (SEQ ID NO:413) present at the finger 1 (F1) position, 15 contained a G in the 5'-most position of the subsite, 15 contained a G in the middle position of the subsite, while, at the 3'-most position of the subsite, 10 contained an A, 3 contained a G and 2 contained a T. Accordingly, this particular amino acid sequence is optimal for binding a GGA triplet from the F1 position.

Please replace the paragraph beginning on page 10, line 10 with the following rewritten paragraph:

Figures 3A, 3B and 3C show site selection data indicating positional dependence of GCA-, GAT- and GGT-binding zinc fingers. The first and fourth (where applicable) rows of each figure show portions of the amino acid sequence of a designed zinc finger protein (F1 column, SEQ ID NOS:220, 1765, 1442, 835 & 1425; F2 column, SEQ ID NOS:220, 1765, 1442, 889, 1425; F3 column, SEQ ID NOS:220, 159, 377, 889 & 1425). Amino acid residues–1 through +6 of each α-helix are listed from left to right. The second and fifth (where applicable) rows show the target sequence, divided into three triplet subsites, one for each finger of the protein shown in the first and fourth (where applicable) rows, respectively. The third and sixth (where applicable) rows show the distribution of nucleotides in the oligonucleotides obtained by site selection with the proteins shown in the first and fourth (where applicable) rows, respectively. Figure 3A shows data for fingers designed to bind GCA; Figure 3B shows data for

A7 cont'd

fingers designed to bind GAT; Figure 3C shows data for fingers designed to bind GGT.

Please replace the paragraph beginning on page 10, line 21 with the following rewritten paragraph:

Figures 4A and 4B show properties of the engineered ZFP EP2C. Figure 4A shows site selection data. The first row provides the amino acid sequences (F3, SEQ ID NO:1100; F2, SEQ ID NO:237; F1, SEQ ID NO:1450) of residues –1 through +6 of the recognition helices for each of the three zinc fingers of the EP2C protein. The second row shows the target sequence (5' to 3'); with the distribution of nucleotides in the oligonucleotides obtained by site selection indicated below the target sequence.

A8

Please replace the paragraph beginning on page 13, line 19 with the following rewritten paragraph:

A D-able subsite within a target site has the motif 5'NNGK3' (SEQ ID NO:4084). A target site containing one or more such motifs is sometimes described as a D-able target site. A zinc finger appropriately designed to bind to a D-able subsite is sometimes referred to as a D-able finger. Likewise a zinc finger protein containing at least one finger designed or selected to bind to a target site including at least one D-able subsite is sometimes referred to as a D-able zinc finger protein.

A9

Please replace the paragraph beginning on page 14, line 16 with the following rewritten paragraph:

Table 6 lists a collection of consensus sequences for zinc fingers and the target sites bound by such sequences. Conventional one letter amino acid codes are used to designate amino acids occupying consensus positions. The symbol "X" designates a nonconsensus position that can in principle be occupied by any amino acid. In most zinc fingers of the C_2H_2 type, binding specificity is principally conferred by residues –1, +2, +3 and +6. Accordingly, consensus

A10

sequence determining binding specificity typically include at least these residues. Consensus sequences are useful for designing zinc fingers to bind to a given target sequence. Residues occupying other positions can be selected based on sequences in Tables 1-5, or other known zinc finger sequences. Alternatively, these positions can be randomized with a plurality of candidate amino acids and screened against one or more target sequences to refine binding specificity or improve binding specificity. In general, the same consensus sequence can be used for design of a zinc finger regardless of the relative position of that finger in a multi-finger zinc finger protein. For example, the sequence RXDNXXR (SEQ ID NO:4060) can be used to design a N-terminal, central or C-terminal finger of three finger protein. However, some consensus sequences are most suitable for designing a zinc finger to occupy a particular position in a multi-finger protein. For example, the consensus sequence RXDHXXQ (SEQ ID NO:4055) is most suitable for designing a C-terminal finger of a three-finger protein.†

A10

Please replace the paragraph beginning on page 32, line 23 with the following rewritten paragraph:

† The amino acid sequence RSDXLXR (SEQ ID NO:4085) (position –1 to +6 of the recognition helix) was found to be optimal for binding to the four GNG triplets, with Asn^{+3} specifying A as the middle nucleotide; His^{+3} specifying G as the middle nucleotide; Ala^{+3} specifying T as the middle nucleotide; and Asp^{+3} specifying cytosine as the middle nucleotide. At the +5 position, Ala, Thr, Ser, and Gln, were tested, and all showed similar specificity profiles by site selection. Interestingly, and in contrast to a previous report (Swirnoff *et al.* (1995) *Mol. Cell. Biol.* **15**:2275-2287), site selection results indicated that three naturally-occurring GCG-binding fingers from zif268 and Sp1, having the amino acid sequences RSDELTR (SEQ ID NO:123), RSDELQR (SEQ ID NO:302), and RSDERKR (SEQ ID NO:1100), were not GCG-specific. Rather, each of these fingers selected almost equal numbers of GCG and GTG sequences. Analysis of binding affinity by gel-shift experiments confirmed that finger 3 of zif268, having the sequence RSDERKR (SEQ ID NO:1100), binds GCG and GTG with

A11

A11 cont'd

approximately equal affinity.

Please replace the paragraph beginning on page 33, line 5 with the following rewritten paragraph:

A12

Based on existing design rules, the amino acid sequence QSGDLTR (SEQ ID NO:220) (-1 through +6) was tested for its ability to bind the GCA triplet from three positions (F1, F2, and F3) within a three-finger ZFP. Figure 3A shows that the QSGDLTR (SEQ ID NO:220) sequence bound preferentially to the GCA triplet subsite from the F2 and F3 positions, but not from F1. In fact, the presence of QSGDLTR (SEQ ID NO:220) at the F1 position of three different three-finger ZFPs resulted predominantly in selection of GCT. Accordingly, an attempt was made to redesign this sequence to obtain specificity for GCA from the F1 position. Since the sequence $Q^{-1}G^{+2}S^{+3}R^{+6}$ (SEQ ID NO:4065) had previously been selected from a randomized F1 library using GCA as target (Rebar *et al.* (1994) *Science* **263**:671-673), a D (asp) to S (ser) change was made at the +3 residue of this finger. The resulting sequence, QSGSLTR (SEQ ID NO:342), was tested for its binding specificity by site selection and found to preferentially bind GCA, from the F1 position, in three different ZFPs (see Figure 2).

Please replace the paragraph beginning on page 33, line 17 with the following rewritten paragraph:

A13

The QSGSLTR (SEQ ID NO:342) zinc finger, optimized for recognition of the GCA subsite from the F1 position, was tested for its selectivity when located at the F2 position. Accordingly, two ZFPs, one containing QSGSLTR (SEQ ID NO:342) at finger 2 and one containing QSGDLTR (SEQ ID NO:220) at finger 2 (both having identical F1 sequences and identical F3 sequences) were tested by site selection. The results indicated that, when used at the F2 position, QSGSLTR (SEQ ID NO:342) bound preferentially to GTA, rather than GCA. Thus, for optimal binding of a GCA triplet subsite from the F1 position, the amino acid sequence QSGSLTR (SEQ ID NO:342) is required; while, for optimal binding of the same subsite

A13 Cont'd

sequence from F2 or F3, QSGDLTR (SEQ ID NO:220) should be used. Accordingly, different zinc finger amino acid sequences may be needed to specify a particular triplet subsite sequence, depending upon the location of the subsite within the target sequence and, hence, upon the position of the finger in the protein.

Please replace the paragraph beginning on page 33, line 29 with the following rewritten paragraph:

Positional effects were also observed for zinc fingers recognizing GAT and GGT subsites. The zinc finger amino acid sequence QSSNLAR (SEQ ID NO:1765) (-1 through +6) is expected to bind to GAT, based on design rules. However, this sequence selected GAT only from the F1 position, and not from the F2 and F3 positions, from which the sequence GAA was preferentially bound (Figure 3B). Similarly, the amino acid sequence QSSHLTR (SEQ ID NO:835) which, based on design rules, should bind GGT, selected GGT at the F1 position, but not at the F2 and F3 positions, from which it preferentially bound GGA (Figure 3C). Conversely, the amino acid sequence TSGHLVR (SEQ ID NO:1425) has previously been disclosed to recognize the triplet GGT, based on its selection from a randomized library of zif268 finger 2. U.S. Patent No. 6,140,081. However, TSGHLVR (SEQ ID NO:1425) was not specific for the GGT subsite when located at the F1 position (Figure 3C). These results indicate that the binding specificity of many fingers is position dependent, and particularly point out that the sequence specificity of a zinc finger selected from a F2 library may be positionally limited.

A14

Please replace the table on page 66 with the following amended table:

TABLE 6

TRIPLET (5'→3')	FINGER (N → C) F1	SEQ ID NO:	F2	SEQ ID NO:	F3	SEQ ID NO:
AGG					RXDHXXQ	4055
ATG					RXDAXXQ	4056
CGG					RXDHXXE	4057
GAA			QXGNXXR	4058		
GAC	DXSNXXR	4059			DXSNXXR	4059
GAG	RXDNXXR	4060	RXSNXXR RXDNXXR	4061 4060	RXDNXXR	4060
GAT	QXSNXXR TXSNXXR TXGNXXR	4062 4063 4064	TXGNXXR	4064		
GCA	QXGSXXR	4065	QXGDXXR	4066		
GCC	EXGTXXR	4067				
GCG	RXDEXXR	4068	RXDEXXR	4068	RXDEXXR RXDTXXK	4068 4069
GCT	QXSDXXR	4070	TXGEXXR QXSDXXR	4071 4070		
GGA			QXGHXXR	4072	QXAHXXR	4073
GGC	DXSHXXR	4074	DXSHXXR	4074		
GGG	RXDHXXR	4075	RXDHXXR	4075	RXDHXXR RXDHXXK	4075 4076
GGT					TXGHXXR	4077
GTA			QXGSXXR QXATXXR	4065 4078		
GTG	RXDAXXR RXDSXXR	4079 4080	RXDAXXR	4079	RXDAXXR	4079
TAG			RXDNXXT	4081		
TCG	RXDDXXK	4082				
TGT			TXDHXXS	4083		

A15

In the claim:

1. (Amended) A method for designing a zinc finger protein comprising a first (F1), a second (F2), and a third (F3) zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus that binds to a target site comprising, in 3' to 5' direction, a first (S1), a second (S2), and a third (S3) target subsite, each target subsite having the nucleotide sequence GNN, the method comprising the steps of

selecting the F1 zinc finger such that it binds to the S1 target subsite, wherein if S1 comprises GAA, F1 comprises the amino acid sequence QRSNLVR (SEQ ID NO:158); if S1 comprises GAG, F1 comprises the amino acid sequence RSDNLAR (SEQ ID NO:130); if S1 comprises GAC, F1 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S1 comprises GAT, F1 comprises the amino acid sequence QSSNLAR (SEQ ID NO:1765); if S1 comprises GGA, F1 comprises the amino acid sequence QSGHLAR (SEQ ID NO:413); if S1 comprises GGG, F1 comprises the amino acid sequence RSDHLAR (SEQ ID NO:127); if S1 comprises GGC, F1 comprises the amino acid sequence DRSHLTR (SEQ ID NO:1506); if S1 comprises GGT, F1 comprises the amino acid sequence QSSHLTR (SEQ ID NO:835); if S1 comprises GCA, F1 comprises QSGSLTR (SEQ ID NO:342); if S1 comprises GCG, F1 comprises RSDDLTR (SEQ ID NO:188); if S1 comprises GCC, F1 comprises ERGTLAR (SEQ ID NO:131); if S1 comprises GCT, F1 comprises the amino acid sequence QSSDLTR (SEQ ID NO:1450); if S1 comprises GTA, F1 comprises the amino acid sequence QSGALTR (SEQ ID NO:1398); if S1 comprises GTG, F1 comprises the amino acid sequence RSDALTR (SEQ ID NO:153); if S1 comprises GTC, F1 comprises the amino acid sequence DRSALAR (SEQ ID NO:184);

A16

selecting the F2 zinc finger such that it binds to the S2 target subsite, wherein S2 comprises GAA, F2 comprises the amino acid sequence QSGNLAR (SEQ ID NO:801); if S2 comprises GAG, F2 comprises the amino acid sequence RSDNLAR (SEQ ID NO:130); if S2 comprises GAC, F2 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S2 comprises GAT, F2 comprises the amino acid sequence TSGNLVR (SEQ ID NO:1442); if S2 comprises GGA, F2

comprises the amino acid sequence QSGHLQR (SEQ ID NO:287); if S2 comprises GGG, F2 comprises the amino acid sequence RSDHLSR (SEQ ID NO:229); if S2 comprises GGC, F2 comprises the amino acid sequence DRSHLAR (SEQ ID NO:1092); if S2 comprises GGT, F2 comprises the amino acid sequence TSGHLSR (SEQ ID NO:1201); if S2 comprises GCA, F2 comprises the amino acid sequence QSGDLTR (SEQ ID NO:220); if S2 comprises GCG, F2 comprises the amino acid sequence RSDDLQR (SEQ ID NO:1844); if S2 comprises GCC, F2 comprises the amino acid sequence DRSDLTR (SEQ ID NO:417); if S2 comprises GCT, F2 comprises the amino acid sequence QSSDLTR (SEQ ID NO:1450); if S2 comprises GTA, F2 comprises the amino acid sequence QSGALAR (SEQ ID NO:3339); if S2 comprises GTG, F2 comprises the amino acid sequence RSDALSR (SEQ ID NO:237); if S2 comprises GTC, F2 comprises the amino acid sequence DRSALAR (SEQ ID NO:184); and

A16 Cont'd

selecting the F3 zinc finger such that it binds to the S3 target subsite, wherein if S3 comprises GAA, F3 comprises the amino acid sequence QSGNLAR (SEQ ID NO:801); if S3 comprises GAG, F3 comprises the amino acid sequence RSDNLTR (SEQ ID NO:231); if S3 comprises GAC, F3 comprises the amino acid sequence DRSNLTR (SEQ ID NO:395); if S3 comprises GAT, F3 comprises the amino acid sequence TSANLSR (SEQ ID NO:377); if S3 comprises GGA, F3 comprises the amino acid sequence QSGHLQR (SEQ ID NO:287); if S3 comprises GGG, F3 comprises RSDHLSR (SEQ ID NO:229); if S3 comprises GGT, F3 comprises the amino acid sequence TSGHLVR (SEQ ID NO:1425); if S3 comprises GCA, F3 comprises the amino acid sequence QSGDLTR (SEQ ID NO:220); if S3 comprises GCG, F3 comprises the amino acid sequence RSDDLTR (SEQ ID NO:188); if S3 comprises GCC, F3 comprises the amino acid sequence DRSDLTR (SEQ ID NO:417); if S3 comprises GCT, F3 comprises the amino acid sequence QSSDLQR (SEQ ID NO:132); if S3 comprises GTG, F3 comprises RSDALTR (SEQ ID NO:153); and if S3 comprises GTC, F3 comprises the amino acid sequence DRSALAR (SEQ ID NO:184);

thereby designing a zinc finger protein that binds to a target site.

Attached hereto is a marked-up version of the changes made to the specification and claims by the current amendment. The attached page is captioned **"Version with markings to show changes made."**

II. REMARKS

The specification has been amended to insert the sequence identification numbers. Claim 1 is pending and has been amended herein to move a clause from the first step into the second step and to correct a typographical error. Support for the amendment can be found throughout the original claim and specification as filed, for example in the Figures. No new matter has been added as a result of this amendment and entry thereof is respectfully requested.

III. CONCLUSION

Applicants believe the claim is in condition for allowance and request early notification to that effect.

Respectfully submitted,

Date: Feb 5, 2002

By: DPasternak

Dahna S. Pasternak
Registration No. 41,411
Attorney for Applicants

ROBINS & PASTERNAK LLP
90 Middlefield Road
Suite 200
Menlo Park, CA 94025
Tel. 650-325-7812
Fax. 650-325-7823